

CLASSIFICAÇÃO DE PADRÕES DE ECG UTILIZANDO REGRESSÃO LINEAR MÚLTIPLA

Arlindo R. Galvão Filho¹

¹ Instituto Tecnológico de Aeronáutica
Dep. Engenharia Eletrônica e de Computação
São José dos Campos - SP

Resumo Este artigo propõe um classificador para detecção de extra-sístoles ventriculares utilizando a regressão linear múltipla. Os dados de ECG MIT-BIH foram utilizados como estudo de caso. Os resultados mostram 97% de acerto do classificador, sendo que os erros ocorridos se devem ao fato da análise dos sinais em questão serem correspondentes a outro canal de ECG.

Palavras Chaves: Classificação de sinais ECG, Regressão Linear Múltipla.

Abstract: This paper proposes a classifier for detecting ventricular extrasystoles using multiple linear regression. The MIT-BIH database were used as case study. The results show 97% accuracy of the classifier, and the errors that occur are because the analysis of the signs in question are related to another channel of ECG.

Keywords: Classification of ECG signals, Multiple Linear Regression.

1 INTRODUÇÃO

Extra-sístole ventricular, ou contração ventricular prematura, é o disparo prematuro do estímulo de contração do ventrículo. Este comportamento é comum em indivíduos saudáveis, porém em diversas cardiopatias, como infarto, estenose aórtica, febre reumática ou insuficiência cardíaca de qualquer origem, podem causar estas disfunções. Nestes casos há possibilidade de piora do quadro clínico do paciente levando a fibrilação ventricular ou até óbito súbito [1].

As extras-sístoles ventriculares possuem características ou morfologias que podem ser observadas no sinal de Eletrocardiograma (ECG). Elas podem aparecer, dentre várias outras formas, com o seguimento ST e onda T com direções opostas ao QRS, ou com o QRS prematuro e alargado (forma bizarra). A identificação desses padrões é de suma importância para uma avaliação sintomática do quadro clínico do paciente. Em alguns casos, os tratamentos podem ser corrigidos por medicamentos, mas existem casos em que é necessário neutralizar a região de ativação do estímulo. Para isto é fundamental a detecção desta região.

Existem casos em que a detecção destes batimentos ectópicos só é possível através da verificação do sinal de ECG por um longo período de tempo, onde o uso do *holter* é indicado. A

verificação destes “sinais alterados” no ECG é feita manualmente por um médico, o que gera um grande esforço e riscos de possíveis erros.

Este trabalho propõe uma forma automatizada de classificar os batimentos normais dos extra-sistólicos através da Regressão Linear Múltipla (RLM); podendo oferecer auxílio de uma análise adicional ao médico.

2 A REGRESSÃO LINEAR MÚLTIPLA

A RLM é empregada quando é necessário estabelecer uma relação envolvendo mais de uma variável independente (x) e a variável de interesse (y)

$$y = x_1 b_1 + x_2 b_2 + \dots + x_n b_n + K \quad (2.1)$$

Dispondo-se de um conjunto de dados de calibração X_{cal} e y_{cal} onde, a melhor estimativa no sentido de mínimos quadrados para o vetor de coeficientes b é dado pela pseudo-inversa de Moore-Penrose [2]:

$$\hat{b} = (X_{cal}^T X_{cal})^{-1} X_{cal}^T y_{cal} \quad (2.2)$$

Em posse do coeficiente de regressão (\hat{b}) é possível fazer a classificação em um novo conjunto de testes no intuito de obter a estimação de \hat{y}

$$\hat{y} = X_{teste} * \hat{b} \quad (2.3)$$

3 ESTUDO DE CASO

Como estudo de caso foi utilizado o Arrhythmia Database do MIT-BIH, banco de dados que está acessível no site Physionet [3]. Nesta base de dados constam registros ECG de diversos pacientes, e todos os registros tiveram os batimentos normais e extra-sistólicos identificados por uma equipe médica. Para este trabalho somente o paciente “200” do MIT-BIH será utilizado; ele contém 1438 registros de batimentos cardíacos normais e 700 extra-sístoles ventriculares como mostra a Figura 3.1.

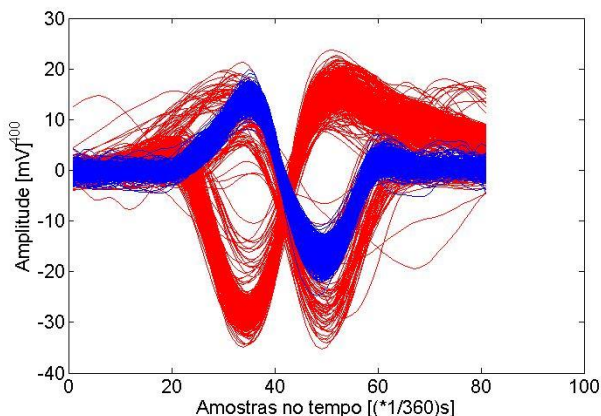


Figura 3.1 – Registro ECG(MLII) referente ao paciente 200.

Cada batimento foi ajustado com 50 amostras, a direita do ponto de marcação do MIT e 50 a esquerda, resultando em 100 unidades. Depois dessa ordenação os sinais foram suavizados com um filtro Savitzky-Golay, com polinômio de segunda ordem e janela de 21 pontos, no intuito de reduzir variações simétricas da linha de base e possíveis ruídos. Resultando em uma redução para 81 unidades, como nota-se na Figura 3.1. O canal de ECG utilizado é o MLII com frequência de amostragem de 360Hz.

Após este pré-processamento, os dados foram divididos aleatoriamente em três conjuntos, como mostra a Tabela 3.1:

É possível observar na Figura 3.1 que existe diversos registros

Tabela 1 – Divisão dos dados

	Batimentos Normais	Extra-sístoles Ventriculares
Calibração	719	350
Validação	360	175
Teste	359	175

extra-sistólicos com características semelhantes aos batimentos normais. Para averiguar melhor o comportamento dos sinais foi aplicada a análise por componentes principais (PCA) como mostra a Figura 3.2.

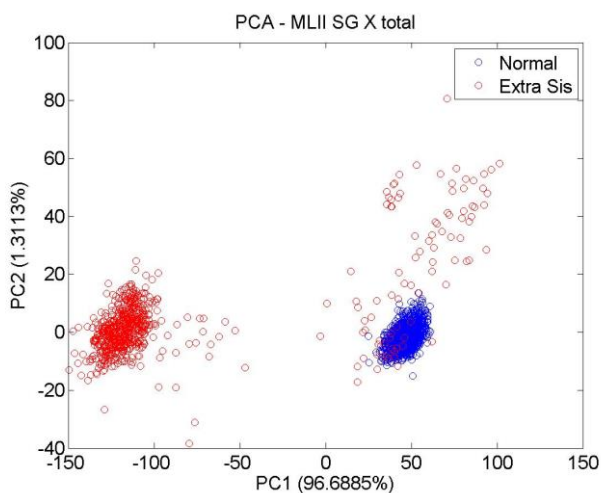


Figura 3.2 – PC1×PC2 dos sinais de ECG total.

Verifica-se que a primeira PC contém a maior variância explicada em relação as outras PC's. Nota-se que existe a separação de dois conjuntos para a maioria dos registros em função da primeira PC; porém há sobreposições de sinais extrasistólicos na região predominantemente normal.

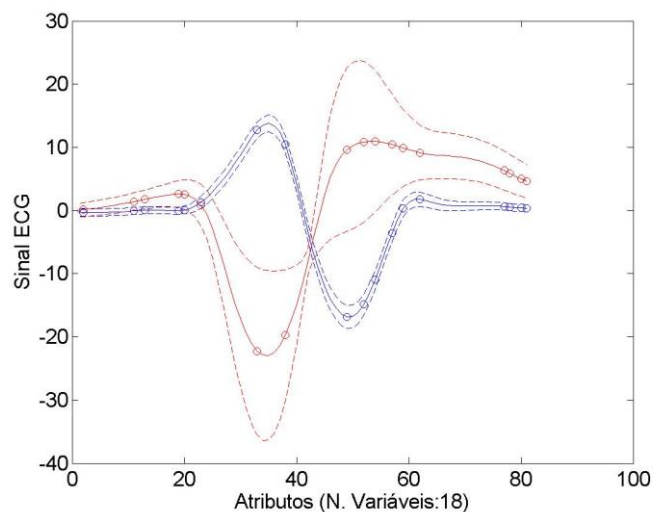


Figura 3.3 – Média dos sinais analisados.

A Figura 3.3 mostra a média e o desvio padrão dos dois conjuntos analisados. Verifica-se que os batimentos normais possuem um desvio padrão pequeno, diferentemente das extrasístoles. É observado também que a região do desvio padrão das extra-sístoles não abrange o padrão dos batimentos normais.

4 RESULTADOS

Ao realizar a classificação utilizando a MLR, houve somente 13 erros respectivos aos sinais extra-sistólicos, como mostra a Figura 4.1

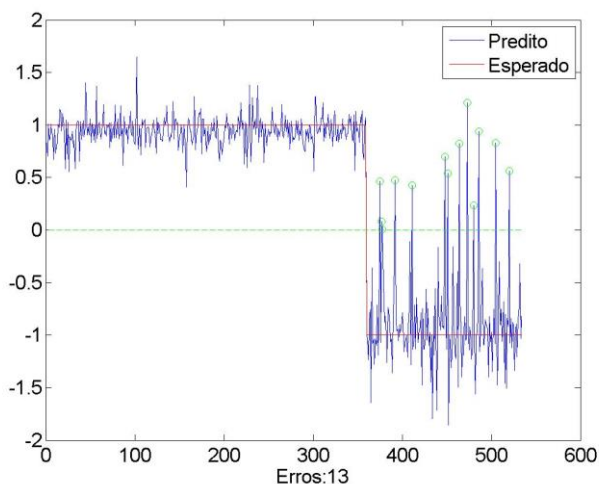


Figura 4.1 – Classificação dos batimentos normais (valores positivos), extra-sistólicos (valores negativos) e os erros (o).

O uso de algoritmos de seleção de variáveis é indicado na MLR pelo problema de colinearidade nos dados. Neste trabalho foi testada a MLR acoplada ao Algoritmo das Projeções Sucessivas (APS) usando somente as amostras selecionadas por ele. O resultado utilizando o APS foi similar à MLR obtendo também 13 erros; mostrando-se desnecessário para este exemplo.

Os sinais respectivos aos erros na classificação são apresentados em verde na Figura 4.2, onde é possível verificar o mesmo padrão dos batimentos normais. Porém, ao analisar os registros ECG sinalizados pelos médicos neste instante de tempo nota-se que a extra-sístole foi detectada pelo canal precordial 1 (V1).

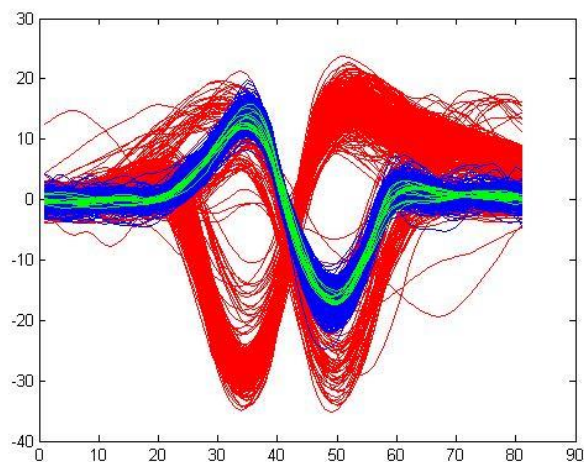


Figura 4.2 – Registro total com os erros.

Desta forma, se somente as informações contidas nos sinais providos do canal MLII forem levadas em conta, o erro do classificador pode ser desconsiderado.

Não foram feitas análises no canal V1, pois existem muitos ruídos e extrapolações talvez provenientes de mal posicionamento dos eletrodos. Os sinais do canal V1 juntamente com os respectivos erros encontrados (em verde) com o modelo anteriormente citado são mostrados na Figura 4.3.

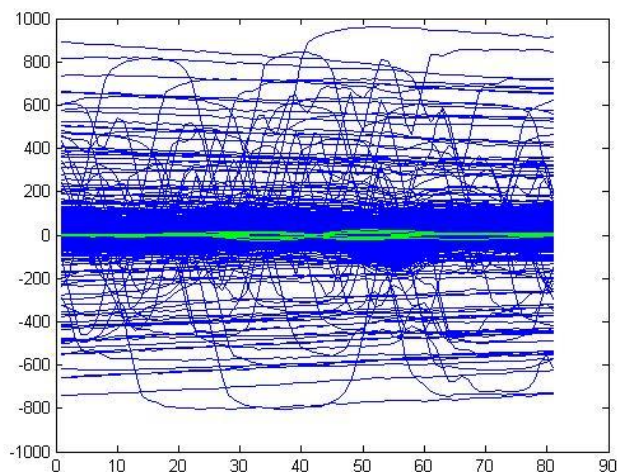


Figura 4.3 – Registro ECG(V1) total com os erros obtidos no canal MLII.

5 CONCLUSÃO

Os resultados mostraram que a classificação das extra-sístoles ventriculares e dos batimentos normais utilizando a MLR obtiveram 97% de acerto, sendo que os 3% restantes podem ter ocorrido devido à análise do canal V1.

REFERÊNCIA BIBLIOGRÁFICA

- Matsumoto, M. M. S.; Lage, D. M. e Galvão, R. K. H. (2008) Agrupamento Morfológico de Sinais Extra-Sístole Ventricular Usando Redes de Kohonen. *XXI Congresso Brasileiro de Engenharia Biomédica*.
- Lawson, C. L. e Hanson, R. J. (1974) Solving Least Square Problems. *Englewood Cliffs: Prentice-Hall*.
- MIT-BIH Arrhythmia Database, Disponível em: www.physionet.org. Acesso em: 08/07/2010.
- Huang, X. and Pan, W. (2003) Linear regression and two-class classification with gene expression data. *Bioinformatic*, Vol. 19 no. 16, pages 2072-207.